Educational research, peer review and editor engagement [*]

Douglas P. Newton, *School of Education, Durham University, UK*

ABSTRACT *Peer review of research articles is known to have shortcomings in a variety of disciplines. Many of these shortcomings are explained by reference to situational, psychological, sociological and ethical influences on reviewers' behaviour. Ways of combating these shortcomings have been proposed, usually aimed at the reviewer. Given that reviewing is voluntary, these proposals tend to be unrealistic. An opportune study indicated that the reviewing of articles for some journals in education showed similar shortcomings. Here, attention is turned largely to the editor and author. Suggestions are made which, in effect, empower the author through the critical engagement of the editor with reviews. For the longer term, it is proposed that professional ethics, sources of bias in reviewing and the practice of reviewing is taught on professional training courses for researchers. Peer review is not confined to articles but is also used to assess other submissions, such as research proposals for grants and the quality of a person's or department's research work. Much can be at stake so the process and outcomes need to be defensible.*

Introduction

Like other academics, those who study education must often teach, research and administrate. But, for advancement, research is often essential and success in it entails publishing that research in academic journals (Snodgrass, 2006). Publication may draw attention to an institution, add to its prestige and earn it income through external assessments of its research activity. In short, a lot hangs on it – unpublished research rarely counts for much (Hopps, 1990).

Whatever the discipline, articles submitted for publication are generally subject to peer review. This is an evaluation of an article by members of the scholarly community to judge its suitability for publication (Michels, 1995; Hames, 2007; Miller & Donati, 2007). According to Burnham (1990), there was a 'casual referring out of articles' in the nineteenth century but peer review did not become institutionalised until the second half of the twentieth century, perhaps to deal with specialisation and an increasing number of submissions. In one sense, the practice works in that it provides journals with articles with a stamp of approval. This does not mean, however, that the stamp is sound. Studies of peer review in education are very rare but those in other fields have shown it can be flawed. For instance, in the USA, Peters and Ceci (1982) re-submitted twelve psychology articles to the journals that had recently published them. Eight of the twelve articles were rejected, mainly for supposed 'methodological flaws'. Garfunkel et al. (1990) took twenty-five reviewed, revised and accepted medical articles and re-submitted them for review to two more referees. The new referees recommended that the articles be revised before

[*] A version of this appears as: Newton, D.P. (2010) Quality and peer review of research: an adjudicating role for editors, *Accountability in Research: Policies and Quality Assurance,* 17(3), 130-145. ISSN 0898 9621

'acceptance'. Bradley (1982) asked authors to comment on the review of their accepted papers. Sixty percent felt that referees had focused on trivia, 40% considered that referee reading had been careless, and 66% felt they were pressured to conform to reviewers' subjective preferences. Editors, of course, can be aware of weaknesses in the reviewing system. For instance, Smith (1999), an editor of the British Medical Journal, concluded that peer review is 'slow, expensive, profligate of academic time, highly subjective, prone to bias, easily abused, poor at detecting gross defects, and almost useless for detecting fraud'. Eysenck and Eysenck (1992), editors themselves, add that the practice delays publication, it is time-consuming and can be costly. One editor commented that 'all who submit articles for publication realize the Monte Carlo nature of the review process' (reported in Eysenck & Eysenck, 1992, p. 394). Why is article submission a gamble?

There are incompetent and negligent reviewers. Incompetence is often illustrated by their comments on statistics. For instance, Baccehetti (2002, p. 1271) showed that reviewers of medical research are guilty of 'finding flaws where there are none'. In particular, reviewers often complain of a 'small' sample size when the statistical analysis is more than capable of producing meaningful results. Ioannidis (2005) points out that, instead, attention should be on the likelihood that findings are true in general. Similarly, reviewers who focus on trivia and minutae and those who are careless might be considered negligent.

There is also bias. For instance, hard-to-read articles tend to be more highly rated (Armstrong, 1980; Sokal, 1996). Armstrong describes this as an instance of the 'Dr Fox Phenomenon'. Dr Fox was an actor whose meaningless exposition of a nonsense topic impressed his audience for its 'clarity' and 'stimulating' content. Another kind of bias favours arguments which agree with the reviewer's beliefs, attitudes and interests and act against those which do not. It can also result in hostility towards new ideas and a preference for orthodoxy (Mahoney, 1977; Armstrong, 1980). On the other hand, well-known, authority figures tend to be favoured (Toulmin, 1972; Armstrong, 1982; Campanario, 1998a). Furthermore, articles, accepted or rejected, are likely to receive five times as many negative comments as positive comments (Bakanic et al., 1987). It looks like reviewers feel they should find fault and look for something to justify rejection, which they most often see as being in the method (Finke, 1990; Tannen, 1998). Eysenck and Eysenck (1992) suggest that reviewers in the West have learned a 'persecution mentality' through experience and perpetuate it themselves. They also point to the way some reviewers make wounding, ad hominen remarks in a 'climate of abuse' which is supported by referees' anonymity (see also Campanario, 1998b, p. 285).

Given these observations, it is not surprising that peer review has been described as a game of chance. In a review of responses to some 3000 natural and social science articles, Cicchetti (1991) found there to be a reliability (inter-reviewer agreement) of only 0.21. Some reviewers agree on their evaluation but disagree on their recommendation. This could stem from, for instance, differences in the judgement of an article's worth. There are also reviewers who reject an article for different reasons. It could be argued that this strengthens the decision to reject but it could equally indicate a fundamental disagreement - what was acceptable to one was not acceptable to the other. Referees' judgements have also been found to lack validity in that they do

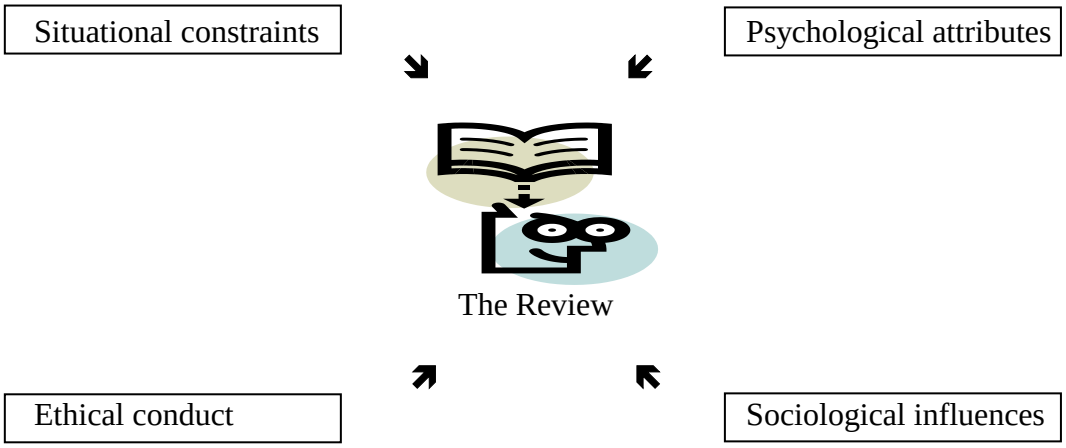
not relate to the subsequent success of articles (Gottfredson, 1978; Eysenck & Eysenck, 1992; Campanario, 1993, 1995, 1996; 1998a).

On this basis, peer review does not always provide much of a stamp of approval. Of course, not all reviewers fall short of the task and some do more than is asked. Cummings et al. (1985) described 'coaches' and 'critics'. Coaches try to be helpful and constructive in their comments. Critics, however, are generally negative and tend to be insensitive to the form and tone of their comments. If an author's work has the misfortune to be reviewed by the latter, it is more likely to be harshly treated and rejected.

Influences on the peer review process

Why do reviewers behave in these ways? These studies suggest it stems from the action and interaction of psychological attributes, sociological influences, situational constraints and ethical conduct, summarised in figure 1.

Figure 1: A model of the peer review process



First, reviewers may be subject to stated and assumed situational constraints. For instance, replications which confirm another's findings are rarely published. This can be a conscious and accepted practice, often justified on the grounds that a journal is oversubscribed and must give preference to more 'important' articles (Hopps, 1990). Non-publication can, of course, distort the record and threatens the validity of subsequent meta-analyses (Moonesinghe et al. 2007). It is much the same for the work of a researcher who proposes, tests and rejects a theory: there is a bias towards positive results.

Second, psychologists could point out that the so-called confirmatory bias is an instance of Heider's assimilation – contrast theory where information concordant with beliefs is accepted and discordant information is rejected (Lindzey & Byrne, 1969). Such behaviour preserves a reviewer's often well-integrated and extensive mental structures which might otherwise have to be radically revised or abandoned. Favouring eminent researchers is an instance of Thorndike's halo effect (Thorndike, 1920). When a research community has agreed that someone is a major contributor to a field, a solitary reviewer may be reluctant to disagree and risk ridicule by rejecting a

new submission. On the other hand, perversely opposing someone's theory or argument is explained by Brehm's reactance theory which describes how some people tend to respond to a view presented to them by adopting the opposite view (Brehms & Brehms, 1966). Such people may be motivated by, for instance, a need to demonstrate some cognitive ability. The Dr Fox Phenomenon has, of course, been known for some time (Naftulin et al., 1973) and could, for instance, reflect a tendency to conceal ignorance and preserve a self-concept. And, of course, psychologists would point out that people vary in the magnitude of these and other traits, knowledge, experience and abilities so some are more conscientious and competent than others.

Third, sociologists could point to the enormous imbalance of power in the reviewer-researcher relationship. As Giddens (2001, p. 420) puts it, 'Power is the ability of individuals or groups to make their own interests or concerns count'. The exercise of this power does not have to be conscious. Prevailing ideologies – the values and beliefs generally held by communities – and what Foucault (1970, 1980) called discourses – ways of thinking about a subject – may be largely unconscious but they can, nevertheless, be powerful forces restricting alternative ways of thinking and speaking (Giddens, 2001, p. 676). Haralambros and Holborn (2000), for instance, describe the way these 'blind members of a society to alternatives [so they] tend to accept the current situation as normal, natural, right and proper'. Lipton (2004, p. 151) has similarly pointed out that 'background beliefs' shape a world view which some find difficult to step outside. On this basis, reviewers reading studies which do not accord with their ideologies and discourses may exercise their power to maintain the status quo by rejecting those studies. This could also help to explain the favouring of the ideas of eminent researchers who, by implication, have already shaped currently accepted discourses. At a shallower level, it could even account for some reviewers' insistence on having an article re-written in their preferred styles. Even a culture of finding fault can be seen as stemming from habituated, unquestioned expectations of what is expected of a reviewer (McIntyre, 1985).

Fourth, the stated purpose of an academic journal is generally to further the general good by making available to the relevant community sound studies which enhance knowledge and understanding in some field. To further that end, reviewers are expected to appraise submissions and make recommendations regarding publication. Researchers who submit articles in good faith could claim a right to have them judged by reference to this general good and not the private good of the reviewer or the protection of the general good of a group to which the reviewer belongs. Readers of the journal could claim a similar right. To the extent that a reviewer's bias is conscious and deliberate, it is a breach of trust which subverts a journal's aims in order to further selfish ends. Many would see this behaviour as unethical (e.g. Kitchener, 1984 whose moral principles include beneficence, nonmaleficence, justice and fairness). We must allow, however, that ideologies and discourses may be unconscious and that bias is not deliberate and is unnoticed by the reviewer. This behaviour could be described as ethically blame-free although it could be argued that the reviewer has a duty to think, to develop competence, to avoid negligence and to question his or her own motives and assumptions (Mabbot, 1966; Foucault, 1970) as the writer and reader are still entitled to an unbiased review. Much of the weakness evident in peer review may not be the result of Machiavellian behaviour but, instead,

could reflect the incompetence, negligence and unconscious bias of reviewers who believe they are behaving ethically.

Does Figure 1 describe the peer review of scholarly work in education? Specific studies of the situation in education are rare and largely uncritical (e.g. Baker, 2002) but, given that educational research has much in common with what has been described here, it seems a reasonable assumption. Nevertheless, even if the assumption is valid, editors may engage with reviews and nullify unwarranted adverse effects. An opportune event provided data relating to the assumption and such remedial action.

Peer review in education

The study of peer review is not easy; editors and reviewers are reluctant to take part and studies which use fake articles waste reviewers' time and risk goodwill (Campanario, 1998a). On the other hand, studies of responses to genuine articles can be open to the criticism that like is not compared with like. Occasionally, when a formal experiment may not be feasible, circumstances present a natural event which can usefully inform discussion.

In a study of the extent to which elementary school books could serve as models of good practice for novice teachers, text was sampled and clauses identified according to rules. Eight articles were written, each on a different school subject, and submitted to subject-oriented, academic journals in education. All articles followed the same pattern: a review of the literature as it related to the subject, a statement of aims, an account of the method, and a discussion of the findings and their implications. The articles had a lot in common, particularly in the method, results and conclusion. Differences were most evident in reference to subject-specific literature, largely in the introductions and discussions. The written reviews and editorial responses provided the data, summarised in Table I.

Two journals accepted and published the articles without forwarding reviews to the authors. Nothing can be said about peer review in these journals. Two reviews were received for each of the other articles. One article (H) was withdrawn from a journal and submitted to another so that two additional reviews were available. Some were photocopies of what appeared to be reviews in their entirety; others were comments selected by the editor. The responses were divided into those relating to generic matters (material more or less common to all articles) and those relating to subject-specific matters. The latter were infrequent and have been omitted from Table I as they could identify of a particular journal. The comments received suggested that the reviewers' attention was most commonly attracted by the introductions and methods.

TABLE I here>

The relatively infrequent subject-specific comments were less useful as differences might reasonably be expected given that the interpretation of the data could vary with the subject. Several reviewers acted as proof readers, even at the detailed level of pointing out a missing bracket. While this may be a useful function (and one that few reviewers seemed able to resist) it says little about the substance of an article and so is not included here.

A natural event allows research where it might otherwise be unwelcome but it has weaknesses. In particular, it is rare for variables to be well-controlled. It could be argued, for instance, that these journals had different standards. At the same time, the articles had much in common but were not identical, particularly when discussing subject-specific matters. Further, editors may have withheld some comments. Similarly, editors' responses may not tell all: there may be more going on behind the scenes. Even allowing that, the comments listed here point to shortcomings like some of those found elsewhere. Some might say these shortcomings are irrelevant as the articles were published. Changes were made which probably improved the articles but changes were also made simply to accommodate reviewers' personal preferences, negligent reading and incompetence. Less experienced authors may have had less success.

The reviews

First, there was the matter of the title. This was variously described as 'motivating' (G), 'eccentric' (H), in need of 'change to reflect the content better' (D) and 'cumbersome' (H′). It could, of course, be all of these but most reviewers, presumably, found it acceptable.

Regarding the introduction, three responses (C, D, G) made the point that it needed to recognise an international audience and, as the articles did not do so, this is certainly a valid criticism but not one that the majority identified. An explanation of textual support for understanding was central to the interpretation of the data. Some responses (for example, E, G) found it to be inadequate. Presumably it was adequate for the majority, one describing it as 'excellent' (H′). Contradictions were also evident in responses from a given journal. For instance, one reviewer found the introduction to be too long while another (G) wanted it to be longer. There was also direct expression of emotion (e.g. 'annoying') and, at best, personal, stylistic preferences (e.g. a dislike of what was perceived to be an 'American' style (D)). One reviewer wished to impose his/her beliefs and preferences on the article in a major way. Being inimical to teachers using textbooks, s/he felt the article might encourage textbook use so wanted the article cast in a form which discouraged their use (D).

The method similarly attracted significant attention, some contradictory and some incompetent. For instance, a reviewer could describe it as 'clear' or 'detailed and clearly written' and, referring to the same article, another reviewer wanted more explanation (D). One reviewer (of G) mistook the method for another approach and insisted on the inclusion of completely inappropriate elaboration and citations. There was also variation in the overall conclusions. A reviewer (of D) wrote that s/he 'looked forward to reading a revised version' but another felt the article was 'very nearly publishable as it stands'. And one reviewer (of F) expressed a dislike of 'facts and figures' and wanted them (that is, the results) removing from the article to make it more readable.

Taken together, some comments undoubtedly pointed to aspects of the articles which needed attention, such as allowing for an international audience. But there is evidence of wide variation in reviewers' responses both to the same aspect and in the aspects they focused on, even in one article. There was evidence of pressing for personal

beliefs and preferences without reasoned justification and of emotive intrusions. There was also evidence of incompetence and inadequacy and of unsubstantiated resistance to ideas not shared by the reviewer. If these were removed from some reviews, there would be little left.

The editors' comments

The editors' comments showed as much variation as those of the reviewers. Some (G, H) appeared to have a very mechanical view of their role: reviews are collected, forwarded to authors and recommendations acted on. Contradictory statements were ignored, as in G, when one reviewer wanted a shorter introduction and considered the method to be clear while the other wanted a longer one and more explanation in the method. While an editor may be unfamiliar with particular research methods, a reviewer's embarrassing error in this respect (G) should have been evident in the contrast between the reviews but it was, nevertheless, presented for action. Other editors seemed to be fairly mechanistic (for example, E) but this may be understandable given that the revisions required were 'minor'. One (H′) responded to an enquiry regarding the need to comply with a reviewer's comment and agreed it could be set aside and another (C) decided that the revisions were optional. F had a similar response to one referee's dislike of 'facts and figures'. There was evidence of one (D) engaging with the reviews critically. This editor's marginal comments on the reviews indicated that s/he had noted the 'American' reference and the instance of confirmatory bias (in this instance, the other side of the coin when discordant information regarding book use was rejected) and both were dismissed. Taken together, there was evidence of a range of responses to the reviews from a mechanical processing of paper to a thoughtful and critical engagement with content. The former approach has the potential to present several kinds of difficulty for authors, not least being rejection without good cause. Equally, it could lead to the unwarranted acceptance of an article. Assuming outright rejection is not the case, there remains the problem of how to deal with contradictory statements, incompetence, expected compliance with reviewers' preferences and background beliefs, emotive responses, prejudice and bias. The latter approach, a critical engagement with reviews, stands a chance of filtering out inadequate comments to arrive at reasonable expectations and a balanced decision – assuming, of course, that the editor is aware of the weaknesses of peer review.

This offers evidence that the model in Figure 1 has relevance for peer review in education. In addition, some editors may work in ways which are unlikely to counter the adverse effects of biased reviews. On this basis, further consideration of the process is justified.

Improving peer review

Suggestions for improving peer review include a code of conduct, guidelines and structured response sheets (Epstein, 1995; Hadjistavropoulos & Bieling, 2000; Hauser & Fehr, 2007; Patterson, 2007; Rojewski & Domenico, 2004). These have merit in that they, at least, indicate what is to count as 'good' behaviour (Austin et al., 1990). From a postmodern perspective, of course, what counts as good behaviour depends on your ideology – there are no absolutes. Even the expectation that criticism should be of the content and not the person is not universally accepted (see, for instance, Hames, 2007). Nevertheless, the restricted community of educational researchers in a field

should agree a code of conduct for its reviewers even if this means that the code varies with field and location. Given the aspirations of many journals to be 'international', such a code or codes would need to accommodate a variety of cultures. As what is perceived to be good professional practice may change with time, the codes would need to be subject to revision. A code, however, has limitations. While it could help those who would be ethical, it is difficult to enforce, difficult to know when it has been breached, it does not prevent incompetence, carelessness and negligence, and a plurality of codes could confuse (Austin et al., 1990). Nor can a code address the unconscious bias of those who see themselves as unbiased. These limitations matter because of the imbalance of power between the reviewer and author.

Some suggestions may only compound the problem. Jayasinghe et al. (2001, 2006), in a study of the peer review of research funding proposals which, in part, have a similar form to an article, found that increasing the number of reviewers and having them review more often makes their recommendations more reliable (see also Marsh et al., 2008). Neff and Olden (2006) suggest a formal assessment by three or four reviewers with a three out of three or four out of four 'decision' rule used for acceptance. Increased reliability and vote counting may make the decision easier, even mechanical, but they do not necessarily improve its quality: reviewing could be reliably biased. Some suggestions could alter the balance of power. Patterson (2007), for instance, argues that enabling discussion between authors and reviewers is more productive than keeping them apart. Conceivably, discussion could lead a reviewer to take another perspective. Open peer review in which the reviewer's identity is revealed has also been found to make reviewers more objective and thoughtful (Armstrong, 1982; Hadjistravropoulos & Beiling, 2000). How practical these are, given that reviewing is voluntary, is open to question. More radical is Armstrong's suggestion that formal reviewing could be eliminated altogether in a law-court approach in which a case is presented, cross-examined and defended on the Internet (Armstrong, 1997). This assumes a willingness of the research community to engage in the prosecution and defence of what could be an interminable case.

Amongst such suggestions, one powerful figure who is generally ignored is the editor. Some editors seem to prefer a mechanical role but this misses the opportunity to alter the power relationship between the reviewer and the author. Editors who are aware of the weaknesses of peer review, engage with reviews, and reflect reasonably on authors' responses to them strengthen the author's hand a little and protect the author's (and reader's) right to a quality control process in which effort has been made to reduce its shortcomings. This means the author may draw an editor's attention to perceived incompetence, negligence, carelessness, unethical behaviour and unconscious bias. Often, an editor will be able to consider such points directly but, on occasions, may wish to seek additional opinions from an editorial board.

Another person who is usually ignored is the author. Given the spectre of author litigation, perhaps on grounds of financial loss, libel or bias and an argument for accountability for failure (Robergs, 2003; Kalles, 2005), it would be wise to allow authors the right to appeal against a decision when they can make a non-trivial, *prima facie* case. Moreover, given the reviewer-author power imbalance, a right of appeal offers some redress to the author. But authors themselves need not be only passive recipients of whatever empowerment is allowed them. Foucault (1988) famously

advised that the unreasonable exercise of power could be resisted by a refusal to comply. In practice, the need to publish is a big stick driving the researcher to accept whatever demands a reviewer makes. Nevertheless, authors may benefit if they stiffened their resolve to make a reasoned objection to an editor.

Finally, to return again to the reviewer but to look to the future, ethics - 'those morally permissible standards of conduct expected of a community' (Mabbot, 1966, p. 45) - need to be a part of the professional training of the educational researcher. The novice researcher also needs to practise reviewing in accordance with one or more codes of conduct and should be made aware of the possibility of conscious and unconscious bias and their sources. Such activity would complement or extend the expertise we expect but it is likely to benefit from structure and should be informed by research. It could also be that new editors would also welcome guidance on developing codes of conduct and ethical practice, on managing appeals, on lending some power to authors, and on engaging with reviews to eliminate evident bad practice. Such programmes are not common but that of the World Association of Medical Editors discusses and produces papers which explore and define responses to matters involved in reviewing (see www.wame.org).

Conclusion

The peer review of educational research can, on occasions, be deficient. Proposals for improving the process usually ignore the voluntary nature of reviewing and so are likely to be relatively ineffective. A model of the process based on situational, psychological, sociological and ethical influences can explain significant deficiencies and, at the same time, point to actions which could lessen their adverse effects. It also draws attention to the other participants in the process who are usually ignored, namely, the editor and the author. For instance, the editor can alter the author-reviewer power relationship to lend more power to the author but, to be effective, editors must acknowledge the potential fallibility of reviewers, engage with reviews and be willing to negotiate with authors. A system which allows a right of appeal for authors could also lend them some power in the relationship. In the longer term, reviewing may be improved through a professional training of educational researchers which gives attention to ethics, professional codes of behaviour, research studies and practice in reviewing.

While articles for publication in research journals have been at the focus of the discussion, peer review also enters into other areas of academic work, as in the assessment of applications for research funds, the public grading of a university's research output and, at times, the assessment of an applicant for a particular post. The reader's attention is drawn to these as other areas where what has been discussed here is likely to have relevance.

Postscript

Seven editors of journals similar to those in the study were asked to comment on some of the suggestions regarding their role. Four responded. They agreed that editors should make the decision about publication, not reviewers. One expressed the view that this decision should rest on reviewers' comments and a personal reading of the article. Another wrote that the process is 'evidence-informed'. But it was interesting that three saw this mainly in terms of the need to resolve reviewer disagreement. The

possibility that reviews may agree *and* be inadequate was not mentioned. There was agreement that editing a journal takes time and two said they did much of the work over weekends. Asked if sharing the load with additional editors would ease engagement, two felt it might take more time because of the need to liaise and the possibility of disagreement, one already used this system and found it was very helpful, and the other saw it as potentially helpful. Two pointed out that editorial boards and 'assistant' editors are sources of advice. Regarding training for the position, most had experienced a mentoring or apprenticeship system where a more experienced person guided the new editor. That this could have weaknesses and insufficiencies was not considered. Nevertheless, the general feeling was that more formal training could be useful. There was some hesitancy about an appeals system, largely because of the time it might take. One editor's response to an appeal is to arrange a further review, another saw it as something for the editorial board to consider. The others took the view that the editor's decision is final.

REFERENCES

Armstrong, J.S. 1980. Unintelligible management research and academic prestige, *Interfaces* 10(2): 80-86.

Armstrong, J.S. 1982. Research on scientific journals: implications for editors and authors. *Journal of Forecasting* 1: 83-104.

Armstrong, J.S. 1997. Peer review for journals: evidence on quality control, fairness, and innovation. *Science and Engineering Ethics* 3: 63-84.

Austin, K.M., Moline, M.E. & Williams, G.T. 1990. *Confronting malpractice*. Newbury Park: Sage.

Bacchetti, P. 2002. Peer review of statistics in medical research: the other problem. *British Medical Journal* 324: 1271-1273.

Bakanic, V., McPhail, C. & Simon, R.J. (1987) The manuscript review and decision-making process. *American Sociological Review* 52: 631-642.

Baker, D. 2002. The peer review process in science education. *Research in Science* Education 32: 171-180.

Bradley, J.V. 1982. Editorial overkill. *Bulletin of the Psychonomic Society*: 19: 271-274.

Brehms, S.S. & Brehms, J.W. 1966. *Pyschological Reactance: A Theory of Freedom and Control*. New York: Academic Press.

Burnahm, J.C. 1990. The evolution of editorial peer review. *The Journal of the American Medical Association* 263(10): 1323-1329.

Campanario, J.M. 1993. Consolation for the scientist: sometimes it is hard to publish papers that are later highly-cited. *Social Studies of Science* 23(2): 342-362.

Campanario, J.M. 1995. Commentary. On influential books and journal articles initially rejected because of negative referees' evaluations. *Science Communication* 16: 304-325.

Campanario, J.M. 1996. Have referees rejected some of the most-cited articles of all times? *Journal of the American Society for Information Science*: 47(4): 302-310.

Campanario, J.M. 1998a. Peer review for journals as it stands today – Part 1. *Science Communication* 19(3): 181-211.

Campanario, J.M. 1998b. Peer review for journals as it stands today – Part 2. *Science Communication* 19(4): 277-306.
Cicchetti, D. 1991. The reliability of peer review for manuscript and grant submissions: a cross disciplinary investigation. *Behavioral and Brain Sciences*, 14: 119-135.
Cummings, L.L., Frost, P.J. & Vakil, T.F. 1985. The manuscript review process: a view from the inside on coaches, critics and special cases. In: *Publishing in the Organisational Sciences*, 2nd edition. Eds L.L. Cummings and P.J. Frost. Homewood: Irwin.
Davis, M. 1999. *Ethics and the university*. London: Routledge.
Epstein, S. 1995. What can be done to improve the journal review process? *American Psychologist* 50: 883-885.
Eysenk, H.J. & Eysenck, S.B.G. 1992. Peer review: advice to referees and contributors. *Personality and Individual Differences* 13: 393-399.
Finke, R.A. 1990. Recommendations for contemporary editorial practices. *American Psychologist* 45: 669-670.
Foucault, M. 1970. *The order of things: an archaeology of the human sciences*. London: Tavistock.
Foucault, M. 1980 *Power and knowledge*. Brighton: Harvester Press.
Foucault, M. 1988. The ethics of care for the self as a practice of freedom. In *The final Foucault*, eds J. Bernauer and D. Rassmussen. Boston: MIT Press.
Garfunkel, J.M., Ulshen, M.H., Hamrick, H.J. & Lawson, E.E. 1990. Problems identified by secondary review of accepted manuscripts. *Journal of the American Medical Association* 263: 1369-1371.
Giddens, A. 2001. *Sociology*, 4th edition. Cambridge: Polity Press.
Gottfredson, S.D. 1978. Evaluating psychological reports: dimensions, reliability, and correlates of quality judgment. *American Psychologist* 33: 920-934.
Hadjistavropoulos, T. & Bieling, P.J. 2000. When reviews attack: ethics, free speech, and the peer review process. *Canadian Psychology* 41(3): 152-159.
Hames, I. 2007. *Peer review and manuscript management of scientific journals: guidelines for good practice.* Malden: Blackwell.
Haralambros, M. & Holborn, M. 2000. *Sociology*, 5th edition. London: Collins.
Hauser, M. & Fehr, E. 2007. An incentive solution to the peer review problem. *PLoS Biology* 5(4): e107, doi:10.1371/journal.pbio.0050107.
Hopps, J.G. 1990. Reflections on 'confirmational response bias among social work journals'. *Science, Technology & Human Values* 15(1): 39-45.
Ionnidis, J.P.A. 2005. Why most published research findings are false. *PLoS Medicine,* 2(12): e334, doi:10.1371/journal.pmed.0020334.
Jayasinghe, U.W., Marsh, H.W., and Bond, N. 2001. Peer review in the funding of research in Higher Education: the Australian experience. *Educational Evaluation and Policy Analysis* 23(4): 343-364.
Jayasinghe, U.W., Marsh, H.W., and Bond, N. 2006. A new reader trial approach to peer review in funding research grants: an Australian experiment. *Scientometrics* 69(3): 591-606.
Kalles, D. 2005. Improving professional conduct in publishing. *Computer* 38(10): 114-116.
Kitchener, K. 1984. Intuition, critical evaluation and ethical principles: the foundation for ethical decisions in counselling psychology. *Counselling Psychology* 12(3): 43-57.

Lipton, P. 2004. *Inference to the Best Explanation*, 2nd edition. London: Routledge.
Lindzey, G. and Byrne, D. 1968. Measurement of social choice and interpersonal alternatives. In *Handbook of Social Psychology*, eds G. Lindzey and E. Aronson. Reading, Mass: Addison-Wesley.
Mabbot, J.D. 1966. *An introduction to ethics*. London: Hutchinson.
Mahoney, M.J. 1977. Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research* 1(2): 161-175.
Marsh, H.W., Jayasinghe, U.W., and Bond, N. 2008. Improving the peer review process for grant applications: reliability, validity, bias, and generalizability. *American Psychologist* 63(3): 160-168.
McIntyre, A. 1985. *After virtue: a study in moral theory*, 2nd edition. London: Duckworth.
Michels, R. 1995. Peer review. *The International Journal of Psychoanalysis* 76: 217-212.
Miller, D.R. and Donati, F. 2007. Peer review policies and the Canadian Journal of Anesthesia: an update for authors and readers. *Canadian Journal of Anesthesia* 54: 1-8.
Moonesinghe, R., Khoury, M.J., and Janssens, C.J.W. 2007. Most published research findings are wrong - but a little replication goes a long way. *PLoS Medicine* 4(2): e28, doi: 10.1371/journal.pmed.0040028.
Natfulin, D.H., Ware, J.E., and Donnelly, F.A. 1973. The Doctor Fox lecture: A paradigm of educational seduction. *Journal of Medical Education* 48: 630-635,
Neff, B.D. and Olden, J.D. 2006. Is peer review a game of chance. *BioScience*, 56(4): 333-340.
Patterson, M. 2007. Bringing peer review out of the shadows. *Public Library of Science* online at http://www.plos.org.
Peters, D.P. and Ceci, S.J. 1982. Peer-review practices of psychological journals: the fate of published articles, submitted again. *Behavioral & Brain Sciences* 5(2): 187-255.
Robergs, R.A. 2003. A critical review of peer review: the need to scrutinize the 'gatekeepers' of research in exercise physiology'. *Journal of Exercise Physiologyonline* 6(2): pp. i-xiii.
Rojewski, J.W. and Domenico, D.M. 2004. The art and politics of peer review. *Journal of Career and Technical Education* 20(2): 1-13.
Smith, R. 1999. Editorial: a beginning that should lead to complete transparency. *British Medical Journal* 318: 4-5.
Snodgrass, R. 2006. Single versus double blind reviewing: an analysis of the literature. *ACM Sigmod Record* 35(3): 8-21.
Sokal, A. 1996. Transgressing the boundaries: towards a transformational hermeneutics of quantum gravity *Social Text* 46-47: 217-252.
Tannen, D. 1998. *The argument culture: moving from debate to dialogue*. New York: Random House.
Thorndike, E.L. 1920. A constant error in psychological rating. *Journal of Applied Psychology* 4: 25-29.
Toulmin, S. 1972. *Human Understanding*. Princeton: Princeton University Press.